

Title	Tensor Factorization for Heterogeneous and Spatio-temporal Data( Abstract_要旨 )
Author(s)	Takeuchi, Koh
Citation	Kyoto University (京都大学)
Issue Date	2019-03-25
URL	<a href="https://doi.org/10.14989/doctor.k21907">https://doi.org/10.14989/doctor.k21907</a>
Right	
Type	Thesis or Dissertation
Textversion	ETD

様式VI

## 博士学位論文調査報告書

論文題目 Tensor Factorization for Heterogeneous and Spatio-temporal Data  
(異種・時空間データ解析のためのテンソル分解法)

申請者氏名 竹内 孝

最終学歴 平成23年3月  
早稲田大学大学院先進理工学研究科電気・情報生命専攻修士課程 修了  
平成31年3月  
京都大学大学院情報学研究科知能情報学専攻博士後期課程  
研究指導認定見込

学識確認 平成 年 月 日 (論文博士のみ)

論文調査委員 京都大学大学院情報学研究科  
(調査委員長) 教授 鹿島 久嗣

論文調査委員 京都大学大学院情報学研究科  
教授 山本 章博

論文調査委員 京都大学大学院情報学研究科  
教授 下平 英寿

( 続紙 1 )

京都大学	博士（情報学）	氏名	竹内 孝
論文題目	Tensor Factorization for Heterogeneous and Spatio-temporal Data （異種・時空間データ解析のためのテンソル分解法）		
（論文内容の要旨）			
<p>本論文は、実世界で現れる異種混合的あるいは時空間的なつながりをもつ複雑な関係データを解析し、高精度での予測や有用な知見を得るための、テンソル分解にもとづく機械学習法についての研究結果をまとめたものであり、全7章から構成されている。</p> <p>第1章は序論であり、本研究の目的とその内容を概観している。実世界の複雑に入り組むさまざまな事象は関係データとして捉えることができるが、その一般的な表現として、しばしば多次元配列が用いられること、また、これに対する解析手法としてテンソル分解が利用されることを述べたうえで、その課題として、疎で不確かなデータに対処するため、対象に関する事前知識を適切に取り込むことの必要性を指摘している。具体的な課題としては、テンソル分解の定式化に補助情報の取り込みや、時空間的な構造を組み込むこと、あるいはその構造自身をデータから発見することを掲げ、これらを解決するテンソル分解手法の開発を提案している。</p> <p>第2章では、本論文で必要となる、非負値テンソル分解、一般化カルバック・ライブラーダイバージェンス、グラフラプラシアン正則化などの、基本的な技術について平易にまとめている。</p> <p>第3章では、複数の行列として表現されるデータを同時に分析して、解釈性の高い分析を行うための行列分解を提案している。提案手法では、一種のデータ内でのパターンに限定せず、異種データをまたいだパターンを抽出することで、より深いデータ分析を可能としている。また、解釈性の高い分析手法として知られる非負行列分解をもとに、これを複数の行列データの同時分解に拡張し、効率的な推定アルゴリズムを提案するとともに、確率的生成モデルとしての解釈も与えている。</p> <p>第4章では、第3章で提案した複数行列の同時分解をテンソルに拡張することで、複数の3項以上の関係データの解析手法を実現している。第3章と同様、効率的な推定アルゴリズムと確率的解釈を与えている。大規模な異種混合実データを用いた実験では、予測精度の向上だけでなく、提案手法がよりコンパクトな表現を得ることで、結果の解釈が容易であることを示している。</p> <p>第5章では、センサーネットワークによって収集される交通状況データなどの、時空間的な広がりを持つデータを分析するためのテンソル分解手法を提案している。センサーネットワークでは、センサー自身のトラブルやネットワークの大域制限などを理由に多くの欠損値がみられ、データの疎性が予測の性能に悪影響を及ぼす。提案手法では、データに仮定できる時間的・空間的な連続性を制約として、非負値テンソル分解に構造正則化の形で組み込むことで、データの疎性を軽減することを実現している。また、最適化手法である交互方向乗数法を用いた効率的なアルゴリズムを提案している。さらに、交通データやバイクシェアリングの実データを用いた評価を行い、提案手法によって予測精度が向上するだけでなく、意味のある時空間パターンを</p>			

抽出できたことを示している。

第6章では、第5章と同じく時空間的なデータを対象としたテンソル分解を議論しているが、第5章では人間が予め与えた、時空間的な事前知識をモデル推定に補助的に用いていたのに対して、第6章では、時空間的な構造もデータから抽出することを目的としている。また、これまでの章で扱っていた内挿的な予測だけでなく、外挿的な将来予測に取り組んでいる。時間的な依存構造を扱うための既存研究である、時系列データを用いた自己回帰正則化をもとに、これに新たに空間構造も取り込むことによる、時空間正則化法を提案している。第5章で提案した方法と異なり、依存関係に関するパラメータも推定対象となるため、事前知識をデータに基づいて修正する効果が望める。実データを用いた実験では、欠損値の補完精度だけではなく、将来予測に関しても良好な予測精度が得られることを確認している。また、定性的な分析として、人間の時空間的な行動パターンを抽出できることを示している。

第7章は結論であり、本論文で得られた成果を要約している。即ち本論文は、実世界で現れる異種混合的あるいは時空間的なつながりをもつ複雑な関係データを解析し、高精度での予測や解釈性の高い結果を得るために、テンソル分解に対象の事前知識を取り込む手法を具体的に提案し、その有効性を実データを用いて実験的に検討したものである。本論文では最後に、将来の展望として、これまでは主に線形モデルの枠組みで設計されていたモデルを、より複雑で表現力の高い非線形モデルに拡張すること、その一方で実データの分析において重要な解釈性を担保するという、二つの側面を両立させる方法を確立すること、また、時空間データ分析の既存の知見を一層モデルに取り込むことを挙げ、本論文を結んでいる。

注) 論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。

論文内容の要旨を英語で記入する場合は、400～1,100 wordsで作成し  
審査結果の要旨は日本語500～2,000字程度で作成すること。

(続紙 2)

(論文審査の結果の要旨)

本論文は、実世界で現れる異種混合あるいは時空間的なつながりをもつ複雑な関係データを解析し、高精度での予測や有用な知見を得るための、テンソル分解にもとづく機械学習法についての研究に取り組んだものであり、得られた主な成果は次の通りである。

1. 行列あるいはテンソルで表現されるデータはしばしば疎性の高いものになるため、これらを高精度で安定して解析することは困難であった。この問題に対して、複数の行列あるいはテンソルを補助情報として用い、これらと同時に分解を行うことで、データの疎性に対処する方法を提案した。また、効率的な推定アルゴリズムを提案するとともに、確率的生成モデルとしての解釈も与えた。さらに、大規模な異種混合実データを用いた実験により、予測精度の向上だけでなく、提案手法がよりコンパクトな表現を獲得することと、結果の解釈が容易であることを確認した。

2. 前述の問題に対する異なる方向からのアプローチとして、データが時空間的な関係をもつ場合にその構造を事前知識として積極的に利用することでデータの疎性に対して有効なテンソル分解手法を提案した。データに仮定できる時間的・空間的な連続性を制約として、非負値テンソル分解に構造正則化の形で組み込んだ最適化問題の定式化と、これを解くための交互方向乗数法を用いた効率的なアルゴリズムを提案した。また、実データを用いた評価を行い、提案手法によって予測精度が向上するだけでなく、意味のある時空間パターンを抽出できたことを確認した。

3. 前述の時空間データの解析手法のさらなる発展として、時空間的な構造もデータから抽出でき、また、内挿的な予測だけでなく、外挿的な将来予測にも利用できるテンソル分解の手法を提案した。時系列データを用いた自己回帰正則化を発展させた、新たな時空間正則化法に基づく最適化問題の定式化と、これを逐次的に解くアルゴリズムを提案した。また、実データを用いた実験では、欠損値の補完精度だけではなく、将来予測に関しても良好な予測精度が得られることを定量的に確認するとともに、人間の時空間的な行動パターンを抽出できることを定性的に示している。

以上、本論文は実世界の疎で不確かさをもった複雑な関係データを高精度で解析するために、異種データや時空間的な構造を取り入れたテンソル分解の方法論の開発に取り組んだ結果をまとめたものであり、学術上・実応用上寄与するところが少なくない。よって、本論文は博士(情報学)の学位論文として価値あるものと認める。また、平成31年2月22日に実施した論文内容とそれに関連した口頭試問の結果、合格と認めた。

注) 論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。

更に、試問の結果の要旨(例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」)を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。

要旨公開可能日： 年 月 日以降